

Can computer aided teaching packages improve clinical care in patients with acute abdominal pain?

F T de Dombal, V Dallos, W A F McAdam

Abstract

Objective—To compare three methods of support for inexperienced staff in their diagnosis and management of patients with acute abdominal pain—namely, with (a) structured data collection forms, (b) real time computer aided decision support, and (c) computer based teaching packages.

Design—Prospective assessment of effects of methods of support on groups of doctors in one urban hospital and one rural hospital.

Setting—Accident and emergency department at Whipps Cross Hospital, London, and surgical wards of Airedale General Hospital, West Yorkshire.

Patients—Consecutive prospective series of all patients presenting to each hospital in specified time periods with acute abdominal pain; total patients in the various periods were 12 506.

Main outcome measures—Diagnostic accuracy of participating doctors, admission rates of patients with non-specific abdominal pain, perforation rates in patients with appendicitis, negative laparotomy rates.

Results—Use of any one modality resulted in improved diagnostic accuracy and decision making performance. Use of structured forms plus computer feedback resulted in better performance than use of forms alone. Use of structured forms plus a computer teaching package gave results at least as good as those with direct feedback by computer.

Conclusions—The results confirm earlier studies in suggesting that the use of computer aided decision support improves diagnostic and decision making performance when dealing with patients suffering from acute abdominal pain. That use of the computer for teaching gave results at least as good as with its use for direct feedback may be highly relevant for those who are apprehensive about the real time use of diagnostic computers in a clinical setting.

Introduction

Over the past two decades various studies have suggested that in some clinical settings, such as patients with acute abdominal pain, the use of computer aided decision support systems may be associated with improved clinical performance as regards diagnostic accuracy and decision making.¹⁻⁶ Yet, despite the favourable results of a large scale study⁷ and the replication of those results in other trials⁸⁻¹⁰ the use of these support systems has not become widespread. One factor contributing to this is the reluctance of some clinicians to introduce computers into real life clinical practice; and several authors have expressed reservations about any real time use of computers in terms of decision support.¹¹

Perhaps the most worrying finding of most previous studies⁷⁻⁸ concerns the “baseline” period—both as regards the low overall accuracy and as regards the learning curve over the six month period which most house surgeons and senior house officers spend on a single firm. This learning curve seems to be flat,

signifying that over the six months the inexperienced doctors in the so called training posts actually learn little if anything at all about dealing with the acute abdomen.

Additional impetus has recently been lent to this problem by an article which (together with subsequent lively correspondence) suggested that among 60 senior house officers in accident and emergency departments most did not receive teaching in general surgery and that what teaching was received did not alter their confidence.¹²

We therefore decided that it would be of interest to offer groups of house surgeons and senior house officers in two NHS hospitals a somewhat different computer aided package from that studied previously. In addition to structured data collection forms (the case for which has been described as “overwhelming”¹³), the doctors were offered computer support in the form of a teaching package, which they were to work through shortly after arriving on the firm in question. Direct access to computer prediction in an individual case was not encouraged and, in practice, was not provided. The question to be answered by the study was therefore simply, would the use of the computer as a teaching device (rather than for direct decision support) be associated with beneficial results similar to those already described?

Subjects and methods

Table I shows the scheme of analysis for the study. In all, 62 doctors took part and data from 12 506 patients were analysed. All patients presented to either Whipps Cross Hospital, London, or Airedale General Hospital, West Yorkshire, suffering from acute undiagnosed pain of less than one week's duration. The criteria for study and methods of analysis were as reported.^{7,8}

TABLE I—Numbers of patients studied at Whipps Cross and Airedale General Hospitals stratified by mode of diagnostic support to senior house officers and house surgeons

Mode of diagnostic support	Whipps Cross	Airedale General	Total
Baseline	1 610	450	2 060
Forms alone	1 589	0	1 589
Forms plus feedback	1 985	5 512	7 497
Forms plus teaching	1 010	350	1 360
Total	6 194	6 312	12 506

In Whipps Cross Hospital a total of 6194 patients were studied in three time periods. These were (a) a baseline period (1982), during which 1610 patients were studied and no assistance was provided to the doctors; (b) a trial period (1982-4), in which 3574 patients were studied and the doctors provided with structured data collection forms plus computer aided decision support in real time; and (c) a teaching period (1989-90; 1010 patients), in which doctors were provided with structured forms and a teaching package.

A similar (but not identical) scheme was followed at Airedale General Hospital, a total of 6312 patients being studied. The three time periods were (a) a

Clinical Information
Science Unit, University of
Leeds, Leeds LS2 9LN
F T de Dombal, FRCS,
director

Accident and Emergency
Department, Whipps Cross
Hospital, London E11
V Dallos, FRCP, consultant in
charge

Airedale General Hospital,
Steeton, West Yorkshire
BD20 6TD
W A F McAdam, FRCS,
consultant surgeon

Correspondence to:
Mr de Dombal.

BMJ 1991;302:1495-7

baseline period (1973-4; 450 patients); (b) an extended trial period (1974-86; 5512 patients), in which doctors were provided with structured forms and given computer aided feedback (but not in real time)*; and (c) a teaching period (1989-90; 350 patients), in which the doctors were given structured forms and a teaching package identical with that at Whipps Cross.

TEACHING PACKAGE

The teaching package aimed at developing those skills in inexperienced doctors which other studies had shown to be central to effective diagnosis and decision making in patients with acute abdominal pain¹⁴—namely, careful and accurate elicitation of symptoms and signs from the patient at presentation; accurate evaluation of these in terms of diagnosis; and, finally, decision making so as to offer the optimal form of treatment to each patient. The teaching package was implemented on an ICL DRS 300 computer and was divided into four separate programs. The programs were aimed at developing separate skills in the doctors concerned, and were as follows:

Definitions—The program on definitions was a simple multiple choice set of questions relating to the definitions of various terms (for example, rebound tenderness, Murphy's sign) employed in the interview and examination of patients with acute abdominal pain.

Patients for practice was a computer held databank of case histories presented in random sequence for the user to diagnose. In all cases the real life outcome was known and was displayed to the user after diagnosis was attempted.

Simulation studies was a series of real emergency situations presented to the user for analysis, comment, and action.

Delphic teaching—In the Delphic teaching program the user compared a personal mental image of a particular disease with a real dataset of 5600 cases from around the world.

For each separate program each student's performance was assessed by using a system of penalties, which were based on the likely impact of the mistake in question on subsequent diagnosis and decision making. The penalties were calculated on the basis of values already established for acute abdominal pain.¹⁵ In each instance experience had shown that final year medical students and house officers could attain a median score of 70%, and a pass mark was therefore set at 80%.

Each program occupied about an hour of the house surgeon or senior house officer's time; and for those able to attain the pass mark at the first attempt the entire package therefore comprised four hours of individual teaching. Failure to attain the pass mark resulted in a request to repeat the format. The average time taken to use the package for each senior house officer or house surgeon was roughly five or six hours in total.

Results

Table II gives the results as regards diagnostic accuracy of doctors in the two hospitals. The baseline accuracy (47% at Whipps Cross, 54% at Airedale General) was in line with other baseline studies in the United Kingdom. When structured forms alone were used an improvement in diagnostic accuracy of around 7% was observed. Use of forms plus computer feedback produced a further improvement to around 65-70% in the two hospitals, irrespective of whether the feedback was direct (Whipps Cross) or indirect (Airedale General). In both hospitals use of structured data collection forms plus the teaching package was associated with a diagnostic accuracy of over 70%, irrespective of the time period studied. Each group of

TABLE II—Percentage diagnostic accuracy of senior house officers and house surgeons during various time periods at Whipps Cross and Airedale General Hospitals using different modes of diagnostic support

Hospital	Mode of diagnostic support			Forms plus teaching
	Baseline	Forms alone	Forms plus feedback	
Whipps Cross	47.1	57.5	65.1	76.0
Airedale General	54.3	0*	66.3	73.1

*Modality of using forms alone not available.

house surgeons or senior house officers achieved an accuracy in this range.

Table III shows comparable findings with respect to decision making, though there were differences between the two hospitals in the mode of admission. At Whipps Cross, where admission was via the emergency department, the proportion of non-surgical cases admitted fell from 28.8% in the baseline (unaided) period to 18.9% during the period when forms (with or without computer feedback) were used, and fell to 10.7% when the mode of forms plus computer teaching was available. At Airedale, where admission was direct to the surgical wards, this comparison was invalid.

TABLE III—Decision making performance of senior house officers and house surgeons at Airedale General Hospital with respect to rates of perforated appendicitis and negative laparotomy and at Whipps Cross Hospital with respect to proportions of non-surgical cases admitted and surgical diagnoses missed. Figures are percentages

	Mode of diagnostic support		
	Baseline	Forms plus feedback	Forms plus teaching
		<i>Airedale General</i>	
Perforated appendicitis	27.0	12.5	11.9
Negative laparotomy	22.0	18.3	15.7
		<i>Whipps Cross</i>	
Cases of non-specific abdominal pain	28.8	18.9	10.7
Surgical diagnoses missed	4.6	2.1	2.4

There the largest fall noted was in perforation rates among patients with appendicitis. In both hospitals, however, there was a pronounced fall in surgical bed night usage, amounting to 13% of all acute surgical bed nights in Airedale and savings of between 65 and 100 bed nights a month at Whipps Cross, when performance in the computer teaching period was compared with baseline values (using methods set out in detail elsewhere^{7,8}).

Statistical analysis of the results is presented in the appendix.

Discussion

Whether doctors and nurses like it or not, the time has arrived for using computers in decision making and patient management¹⁶; and the evidence in support of formally structured patient interview pathways (with or without computer aided diagnostic programs) is now overwhelming, and these should be routine.¹⁷ Yet such support has not been widely adopted—and one reason may well be the understandable reluctance of doctors to adopt computer systems which seem to take over their diagnostic and decision making functions.¹¹

What our studies offer is an alternative which may be more acceptable to many doctors and which may go at least some way towards rectifying difficulties in providing teaching for busy junior staff. Our findings indicate that when doctors were provided with both structured data collection forms and computer teaching their performance was at least as good as that in other trials when the computer was used for decision

support. This observation goes some way towards supporting the hypothesis that in these earlier trials the main effect of the computer was related not so much to the arrival of "artificial intelligence" as to its educational effect and its effect in acting as a stimulus to good clinical practice as defined in advance by a peer group.

Such conclusions must remain tentative. Our studies are clearly imperfect, depending as they do on comparisons between time periods (double blind controlled trials are extremely difficult to conduct in this subject), and much further work remains to be done in determining the educational process by which the effects are achieved.

The use of historical controls must inevitably raise the question whether the results of these studies were biased by a "drift" with time—that is, by a steady improvement in diagnosis which has taken place over the past 20 years. In practice there is now very good evidence from eight studies including 18 hospitals and 8500 patients that this was not the case.^{1,3,5,7,8,10,18}

Indeed, if anything, the reverse was true. The most recent studies in seven hospitals showed a diagnostic accuracy in the late 1980s among unaided doctors of only 40%.¹⁸ If anything, therefore, the use of historical controls biases our study against the teaching package introduced; and certainly there is no evidence from wide studies within the United Kingdom that our results were due to a steady improvement with time.

In summary, therefore, two conclusions are permissible at this stage. Firstly, we agree with other workers¹⁷ that there is absolutely no excuse for current performance levels in dealing with patients suffering from acute abdominal pain. Secondly, for those who are wary of using computers in a decision support mode an alternative is now available which appears equally effective.

We acknowledge the support and help of many colleagues in the two hospitals during the study. Particularly we thank those who collected and analysed data (Jason Mellows, Brenda Brock, Susan Clamp, May Chan, Karen Wardle). Studies in Airedale General Hospital (1974-6) and Whipps Cross Hospital (1982-4) were supported with grants from the DHSS. Computing facilities in 1989-90 in both hospitals were made available by Medical Portfolio Ltd. Finally, and particularly, we thank the senior house officers and house surgeons, whose experience forms the basis of this paper.

Appendix

STATISTICAL ANALYSIS

There has been considerable debate about the best method of statistical analysis of data such as those in this analysis.¹⁹ In particular, attention has been drawn to the need for consistency in overall clinical procedures, in case mix between different groups of patients, and in analysing diagnostic accuracy (a) to distinguish between minor and major diagnostic errors and (b) to analyse data in terms of both overall accuracy and accuracy of each individual doctor.

In this analysis there was no difference in clinical or admission procedures in either hospital during the period

studied. In each hospital the case mix remained similar in all time periods—though there was a consistently higher prevalence of non-surgical patients in the Whipps Cross data, where patients were seen in an accident and emergency department.

In table II (diagnostic performance) the differences between each mode of diagnostic support were significant ($p < 0.01$) by simple χ^2 analysis of the overall accuracy in the groups of patients studied. When non-parametric analyses were applied to the performance of individual doctors there was a significant difference ($p < 0.02$) between each of the first three modes. Nevertheless, the difference between the forms and feedback mode and the forms and teaching mode just failed to reach significance ($p \approx 0.07$). None of the differences between the two individual hospitals were significant.

As regards decision making performance (table III), in Airedale General Hospital the fall in perforation rates was significant in both trial periods compared with baseline ($p < 0.01$); the fall in negative laparotomy rates was not significant. At Whipps Cross Hospital the proportion of non-surgical patients admitted was significantly lower in both study periods compared with baseline ($p < 0.01$), but the difference between the two study periods just failed to reach significance ($p \approx 0.06$; Kolmogorov-Smirnov test for individual doctor performance).²⁰ The fall in surgical diagnoses missed was not significant.

- 1 de Dombal FT, Leaper DJ, Staniland JT, Horrocks JC, McCann AP. Computer-aided diagnosis of acute abdominal pain. *BMJ* 1972;ii:9-13.
- 2 Wilson PD, Horrocks JC, Yeung CK, Lyndon PJ, Page RE, de Dombal FT. Computer-aided diagnosis of acute abdominal pain using simple inexpensive equipment and "untrained" personnel. *BMJ* 1975;iii:73-5.
- 3 Gunn AA. The diagnosis of acute abdominal pain with computer analysis. *J R Coll Surg Edinb* 1976;21:170-2.
- 4 McAdam WAF, Davenport P, de Dombal FT. Six years' experience of computer-aided diagnosis in a district general hospital. Report to DHSS. London: HMSO, 1981.
- 5 Boom R. Improvements of internists' diagnostic performances by systematic computer evaluations. In: *Proceedings of Medinfo 80*. Amsterdam: North Holland, 1980:1760.
- 6 de Dombal FT. The OMGE acute abdominal pain survey. Progress report. *Scand J Gastroenterol* 1988;23(suppl 144):35-42.
- 7 Adams ID, Chan M, Clifford PC, et al. Computer aided diagnosis of abdominal pain: a multicentre study. *BMJ* 1986;293:800-4.
- 8 Hancock DM, Heptinstall M, Old JM, et al. Computer-aided diagnosis of acute abdominal pain. The practical impact of a "theoretical exercise." *Theoretical Surgery* 1987;2:99-105.
- 9 Clifford PC, Chan M, Hewett DJ. The acute abdomen: management with microcomputer aid. *Ann R Coll Surg Engl* 1986;68(4):182-4.
- 10 McAdam WAF, Brock BM, Armitage T, Davenport P, de Dombal FT. Twelve years' experience of computer-aided diagnosis in a district general hospital. *Ann R Coll Surg Engl* 1990;72:140-6.
- 11 Balla JI, Elstein AS, Christensen C. Obstacles to acceptance of clinical decision analysis. *BMJ* 1989;298:579-82.
- 12 Morris F, Cope A, Hawes S. Training in accident and emergency: views of senior house officers. *BMJ* 1990;300:165-6.
- 13 Paterson-Brown S, Vipond MN. Modern aids to clinical decision-making in the acute abdomen. *Br J Surg* 1990;77:13-8.
- 14 Leaper DJ, Gill PW, Staniland JR, Horrocks JC, de Dombal FT. Clinical diagnostic process: an analysis. *BMJ* 1973;iii:569-74.
- 15 Staniland JR, Ditchburn JR, de Dombal FT. Clinical presentation of the acute abdomen—a series of 600 cases. *BMJ* 1972;iii:393-8.
- 16 Hewett DJ. The age of information technology. *Hospital Update* 1990;16(4):283.
- 17 Hoffman J, Rasmussen OO. Aids in the diagnosis of acute appendicitis. *Br J Surg* 1989;76:774-9.
- 18 Rowell KV, Annis S. Provision of decision support for clinicians—some lessons in communications from a multi-centre project. *Theoretical Surgery* (in press).
- 19 Schwartz D, Lellough J. Explanatory and pragmatic attitudes in therapeutic trials. *J Chronic Dis* 1967;20:637-48.
- 20 Massey FJ Jr. The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association* 1951;46:68-78.

(Accepted 24 April 1991)

THE MEMOIR CLUB

One day I was sent for by headquarters and stood aside to allow another officer to go up the stairs first. He was given a posting to the Middle East and I was given a relieving job with the 8th Battalion, the Green Howards. I found out afterwards that first come had been first served, and that if I had preceded him I would have been sent to the Middle East and he would have gone to the Green Howards.

The far reaching influence of trivial decisions and actions is, of course, a philosophical commonplace. Yet this particular example of the working of fate impressed me very strongly. But for a mere gesture of politeness I might have become a surgeon or been killed, I might never

have gone to Australia, and would certainly never have met my wife. It is inadvisable to pursue such a train of thought too far, for it leads to questions like: "Who am I?" and "Why am I here?" and after your mind has been occupied by these matters for some time they come for you in a plain van and take you away. But clearly that day my lifeline divided, one fork leading to this autobiography and the other trailing off into the unknown.

From *Not a Proper Doctor* by David Sinclair. Published under the *BMJ's* Memoir Club imprint. ISBN 0 7279 0279 2. Price: Inland £14.95; abroad £18.50. BMA members: Inland £13.95; abroad £17.50.